



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

Journal of Computational and Applied Mathematics 205 (2007) 72–87

JOURNAL OF  
COMPUTATIONAL AND  
APPLIED MATHEMATICS[www.elsevier.com/locate/cam](http://www.elsevier.com/locate/cam)

# A posteriori estimation of the linearization error for strongly monotone nonlinear operators

Alexandra Chaillou<sup>a</sup>, Manil Suri<sup>b,\*</sup><sup>a</sup>Department of Mathematics, College of Notre Dame, Baltimore, MD 21210, USA<sup>b</sup>Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, USA

Received 21 November 2005; received in revised form 12 April 2006

## Abstract

We investigate the a posteriori estimation of the modeling (or linearization) error which arises when a nonlinear problem is replaced by a linear model. Using the context of strongly monotone operators, we construct a computable upper estimator for this error, and also provide an estimator that gives a lower bound. Several numerical results illustrating our theory are provided.

© 2006 Elsevier B.V. All rights reserved.

MSC: 65N15; 65J10; 65J15

Keywords: A posteriori; Estimator; Linearization; Nonlinear elasticity

## 1. Introduction

The investigation of physical phenomena involves several modifications, introduced to simplify one or more aspects of the problem. The primary such simplification occurs in the modeling process, whereby a phenomenon is described by ascribing to it a mathematical model. Such models, which can be quite complicated, are often further simplified to make the problem more computationally tractable. Our concern in this paper is the estimation of the error introduced when a nonlinear model (which we call the *original problem*) is replaced by a linear model. This error is given by

$$e_L = u - u_L. \quad (1.1)$$

Here  $u$  is the *exact solution* of the original problem (i.e., the nonlinear model) and  $u_L$  is the solution of the simplified linear model (the *linearized solution*). This error may be measured in different norms depending upon the goals of the computation (for instance energy norm error, error in the value of  $u$  at a point, etc.).

We call (1.1) the *modeling error*. In practice, a second error is usually introduced when the linearized mathematical model is solved using a numerical method, such as the finite element method. We consider the additional effects of such *discretization errors* in [4]—our goal in this paper is to investigate a posteriori estimation of only the modeling error (1.1), under the assumption that  $u_L$  is calculated exactly. We set our investigation in the context of strongly monotone

\* Corresponding author.

E-mail address: [suri@math.umbc.edu](mailto:suri@math.umbc.edu) (M. Suri).

<sup>1</sup> The work of this author was supported in part by the National Science Foundation under Grant DMS-0074160.

operators, which provide a general abstract framework for our results. These are introduced in Section 2, where we also describe several examples that lead to such operators. Existence and uniqueness for such problems can be obtained if certain monotonicity and continuity conditions hold—these conditions also play an important role in the estimators constructed in subsequent sections. We verify these conditions for the examples introduced, which include elasticity with nonlinear stress–strain relations, the  $p$ -Laplacian, and some variants.

In Section 3, we introduce a linearization of the abstract nonlinear problem and derive estimates for  $e_L$  in appropriate norms  $\|e_L\|$ . We present several computational results showing how our a posteriori error estimators for the modeling error  $e_L$  work in practice. A method of measuring the accuracy of any estimator  $\mathcal{E}$  for an error  $\|e_L\|$  is via its *effectivity index*  $\kappa$  defined by [1,2]

$$\kappa = \frac{\mathcal{E}}{\|e_L\|}. \quad (1.2)$$

As we see from our computational examples ahead, this effectivity index (while very well-behaved for some of our examples) can also be large in some cases (Examples 3 and 4) for isolated values of  $u$  and  $u_L$ . However, the resulting error bounds can still be practically very useful. For instance, in materially nonlinear elasticity problems, there can often be a substantial difference between the actual and assumed values of such quantities as the yield stress. (As an example, a 20% figure is reported for the 5454H32 aluminum alloy in [8], where the yield stress is seen to be extremely sensitive to the thickness of the material.) As a result, even if the effectivity index is 3 or 4 (as in some of our examples), the error due to linearization will still be negligible compared to the error induced by the uncertainty in a quantity such as the yield stress. We mention that in contrast, the modulus of elasticity (which would be the main constant determining the response for the linear problem) does not show such variability.

Let us remark, moreover, that isolated values of  $\kappa$  may not provide a good evaluation of an estimator in any case. A more complete picture in the spirit of asymptotic exactness [1,2] is given by a characterization of the behavior of  $\kappa$  as  $\|e_L\| \rightarrow 0$ . An example of this behavior is illustrated in our Example 4.

We mention a previous work [6], that is relevant in the context of this paper. There, the modeling error for a materially nonlinear problem having a piecewise linear stress–strain relationship was investigated. This corresponds to our Example 1 in Section 2. We obtain similar results to those in [6], although our method of proof is different, being designed to accommodate more general problems. (The goal of [6] was also to consider the effect of singularities in the exact solution, for which results more detailed than the ones here were derived—we do not address this issue.)

Finally, the a posteriori estimators we derive can also be useful in developing a feedback strategy to determine whether the dominant part of the total error is from modeling or discretization. We refer to [3,4] for a discussion and examples.

## 2. The model problem

In this section we describe an abstract formulation of our nonlinear problem (the *exact problem*) and discuss examples which fit into its format.

### 2.1. Abstract formulation

Let  $V$  be a reflexive Banach space and  $V'$  be its dual, with norms denoted by  $\|\cdot\|_V$  and  $\|\cdot\|_{V'}$ , respectively. Let  $A : V \rightarrow V'$  be a nonlinear operator that satisfies the following three conditions (we will use  $Au(v)$  to denote the operator  $Au$  evaluated at  $v$ ).

**M0 :**  $A0 = 0$ .

**M1 :**  $A$  is strongly monotone, i.e., there exists a strictly increasing function  $\chi : [0, \infty) \rightarrow \mathbb{R}$  with

$$\chi(0) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \chi(t) = \infty$$

such that

$$\forall u, v \in V \quad (Au - Av)(u - v) \geq \chi(\|u - v\|_V) \|u - v\|_V. \quad (2.1)$$

**M2:**  $A$  is Lipschitz continuous for bounded arguments, i.e., for any ball,  $B(0; r) = \{v \in V : \|v\|_V \leq r\}$ , there exists a constant  $\Gamma(r)$  such that

$$\forall u, v \in B(0; r) \quad \|Au - Av\|_{V'} \leq \Gamma(r) \|u - v\|_V. \quad (2.2)$$

Then, we have the following theorem.

**Theorem 2.1.** *Let  $A : V \rightarrow V'$  satisfy **M0**–**M2**. Then the problem*

$$Au = F, \quad F \in V' \quad (2.3)$$

*has a unique solution satisfying*

$$\chi(\|u\|_V) \leq \|F\|_{V'}. \quad (2.4)$$

**Remark 2.1.** We have taken condition **M0** for simplicity. In the general case, if  $A0 \neq 0$ , instead of (2.4) we would obtain

$$\chi(\|u\|_V) \leq \|F\|_{V'} + \|A0\|_{V'}.$$

Let us now discuss some notation and terminology we will be using in this paper. We let  $\Omega \subset \mathbb{R}^d$  for  $d = 1, 2$  be a convex bounded open domain with piecewise smooth boundary,  $\Gamma = \partial\Omega$ . We will use the usual Sobolev space notation for spaces  $W^{k,p}(\Omega)$ ,  $W_0^{k,p}(\Omega)$  ( $p > 1$ ). For the case  $p = 2$  we denote  $W^{k,2}(\Omega)$  as  $H^k(\Omega)$ . Our space  $V$  will always be a Sobolev space of the form  $W_0^{1,p}(\Omega)$  with  $p \geq 2$ , though our results can be easily generalized to other cases, such as  $\Omega \subset \mathbb{R}^3$  or more general boundary conditions, for which  $W_0^{1,p}(\Omega) \subset V \subset W^{1,p}(\Omega)$ . While the proofs given here are valid for any norm on the space  $V$  equivalent to  $\|v\|_{W^{1,p}(\Omega)}$ , we will use  $\|v\|_V = |v|_{W^{1,p}(\Omega)}$  (the seminorm is a norm, since we have  $V = W_0^{1,p}(\Omega)$ ). In the sequel, the notation  $\|\cdot\|_{k,p,\Omega}$  shall be used to represent  $\|\cdot\|_{W^{k,p}(\Omega)}$ , with the subscript  $\Omega$  dropped when understood.

Let us turn to the operator  $A : V \rightarrow V'$ . In our examples,  $A$  will always take the form

$$Au(v) = \int_{\Omega} \sigma(x, \nabla u) \cdot \nabla v \, dx \quad (2.5)$$

for  $v \in V$ . Here  $\sigma : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  will be a suitable function satisfying  $\sigma(x, 0) = 0$ . This ensures that  $A$  satisfies **M0**. In the examples we consider here, we will always have

$$\sigma(x, \xi) = a(x, |\xi|)\xi, \quad (2.6)$$

where  $a$  is defined on  $\Omega \times \mathbb{R}^+$  (we will often omit mentioning the  $x$  dependence in both  $\sigma$  and  $a$ ).  $F \in V'$  will be given by

$$F(v) = \int_{\Omega} f v \, dx, \quad (2.7)$$

where  $f \in L_q(\Omega)$  with  $1/p + 1/q = 1$ . This condition can be weakened: for example, if  $p = 2$ , all we need is  $f \in H^{-1}(\Omega)$ .

**Remark 2.2.** The conditions **M1** and **M2** can be weakened. For instance, for the special case (2.6) the paper [7] gives more general conditions on  $a(\cdot, |\xi|)$  that guarantee existence and uniqueness for problem (2.3).

**Remark 2.3.** Although only the scalar case (2.5) is considered here, our results extend readily to systems of equations (such as in the elasticity problem).

## 2.2. Preliminary results

Let us now present some theorems that will be useful for analyzing our examples. We begin with a result that will help establish **M1**.

**Theorem 2.2.** Let  $V = H^1(\Omega)$  and let  $A$  be given by (2.5) such that for any  $x \in \Omega$ ,  $\sigma(x, \cdot) \in C^1(\mathbb{R}^d)$ . Also, let  $D_\eta \sigma(x, \eta)$ , the Jacobian of  $\sigma$ , be uniformly positive definite on  $\Omega \times \mathbb{R}^d$ , i.e.,

$$\xi^T D_\eta \sigma(x, \eta) \xi \geq \alpha |\xi|^2, \quad (2.8)$$

for some constant  $\alpha > 0$  independent of  $x \in \Omega$ ,  $\xi, \eta \in \mathbb{R}^d$ . Then  $A$  satisfies **M1** with

$$\chi(\|u\|) = \alpha \|u\|_V. \quad (2.9)$$

**Proof.** Let  $x \in \Omega$ ,  $\xi_1, \xi_2 \in \mathbb{R}^d$ . Using (2.8) with  $\xi = \xi_1 - \xi_2$  we have

$$\begin{aligned} (\xi_1 - \xi_2)^T (\sigma(x, \xi_1) - \sigma(x, \xi_2)) &= \int_0^1 (\xi_1 - \xi_2)^T D_\eta \sigma(x, \eta) |_{\eta=\xi_2+t(\xi_1-\xi_2)} (\xi_1 - \xi_2) dt \\ &\geq \alpha |\xi_1 - \xi_2|^2. \end{aligned}$$

The result easily follows.  $\square$

**Corollary 2.2.1.** For the special case (2.6) with  $V = H^1(\Omega)$ , **M1** is satisfied with

$$\chi(\|u\|_V) = \alpha \|u\|_V \quad (2.10)$$

provided

$$\min_{x \in \Omega, \xi \in \mathbb{R}^d} a(x, |\xi|) \geq \alpha \quad (2.11)$$

and  $a$  is increasing, i.e., the derivative with respect to  $r$  satisfies

$$a_r(x, r) \geq 0 \quad (x, r) \in \Omega \times \mathbb{R}^+.$$

The conditions of the corollary are sufficient to ensure that  $\sigma$  satisfies (2.8), as can be seen by writing out the Jacobian (see [3] for details). Hence the proof follows by Theorem 2.2.

Next, we present a result for establishing **M2**. We need the following lemma which follows easily from the Minkowski Inequality.

**Lemma 2.3.** For  $A$  as defined by (2.5), with  $V = W_0^{1,p}(\Omega)$ , we have

$$\|Au\|_{V'} \leq \|\sigma(\nabla u)\|_{0,q}, \quad (2.12)$$

$$\|Au - Av\|_{V'} \leq \|\sigma(\nabla u) - \sigma(\nabla v)\|_{0,q}, \quad (2.13)$$

where  $1/p + 1/q = 1$ .

**Remark 2.4.** With different boundary conditions, we can get equality in (2.12) and (2.13). For instance, in one dimension, all we need is for a Dirichlet condition to be imposed only at one end point. Neumann conditions everywhere will also allow us to get equality.

Using (2.13), we easily obtain the following theorem for establishing **M2**.

**Theorem 2.4.** Let  $A$  be defined by (2.5). If  $\sigma$  satisfies

$$\|\sigma(\nabla w) - \sigma(\nabla z)\|_{0,q} \leq \Gamma(r) \|\nabla w - \nabla z\|_{0,p} \quad (2.14)$$

for any  $w, z \in B(0; r)$ , then  $A$  satisfies **M2**.

### 2.3. Examples

Let us now consider our specific examples.

**Example 1 (Bilinear elasticity).** We first consider a one-dimensional example with  $V = H_0^1(\Omega)$ , with  $\Omega = (0, 1)$  (for a higher dimensional analog see [6]). We consider the function

$$\sigma(u_x) = \begin{cases} k(u_x + U_Y) - U_Y, & u_x < -U_Y, \\ u_x, & -U_Y \leq u_x \leq U_Y, \\ k(u_x - U_Y) + U_Y, & U_Y < u_x. \end{cases} \quad (2.15)$$

This represents a bilinear stress–strain relationship. Here  $U_Y \in (0, 1)$  represents the strain corresponding to the yield stress and  $k \in (0, 1)$ . This satisfies the special case (2.6) with  $a(r)$  given by

$$a(r) = \begin{cases} 1, & r \leq U_Y, \\ k + \frac{(1-k)U_Y}{r}, & r > U_Y. \end{cases} \quad (2.16)$$

Let us now verify that conditions **M1** and **M2** are satisfied.

**Condition M1:** To show this condition is satisfied, let, for any  $u, v \in V$ ,

$$\Omega_1^{u,v} = \{x \in \Omega : |u_x|, |v_x| \leq U_Y\}, \quad \Omega_2^{u,v} = \Omega / \Omega_1^{u,v}. \quad (2.17)$$

The Jacobian  $\sigma'$  satisfies

$$\sigma'(\xi) = 1 \quad \text{on } \Omega_1^{u,v}, \quad \sigma'(\xi) \geq k \quad \text{on } \Omega_2^{u,v}.$$

Then the same argument as the proof of Theorem 2.2 gives us that  $A$  satisfies **M1** with

$$\chi(\|u - v\|_V) = k\|u - v\|_V. \quad (2.18)$$

(Note that all we need is for  $\sigma'$  to be integrable, not continuous as in Theorem 2.2.)

**Remark 2.5.** We see, in fact, that

$$(Au - Av)(u - v) \geq \|u - v\|_{1,2,\Omega_1^{u,v}}^2 + k\|u - v\|_{1,2,\Omega_2^{u,v}}^2.$$

When neither  $|u_x|$  nor  $|v_x|$  exceed  $U_Y$  this just reduces to

$$(Au - Av)(u - v) \geq \|u - v\|_{H^1(\Omega)}^2,$$

so that in this case (2.18) can be improved to

$$\chi(\|u - v\|_V) = \|u - v\|_V.$$

**Condition M2:** To show **M2** is satisfied, it is sufficient to consider only the case  $u_x > 0, v_x > 0$ . Then for each of the three cases  $|u_x| < U_Y$  and  $|v_x| < U_Y$ ,  $|u_x| \geq U_Y$  and  $|v_x| < U_Y$ , and  $|u_x| \geq U_Y$  and  $|v_x| \geq U_Y$ , we may easily verify that

$$\|\sigma(u_x) - \sigma(v_x)\|_{0,2} \leq \|u_x - v_x\|_{0,2}. \quad (2.19)$$

By Theorem 2.4, **M2** is satisfied with  $\Gamma(r) = 1$ .

Using Theorem 2.1 we get the following result.

**Theorem 2.5.** Let  $A$  be given by (2.5) with  $\sigma(\nabla u)$  given by the bilinear function described in Example 1 on the space  $V = H_0^1(\Omega)$ . Then  $A$  satisfies **M1** with

$$\chi(\|u\|_V) = k\|u\|_V. \quad (2.20)$$

Moreover,  $A$  satisfies **M2** with  $\Gamma(r) = 1$ . Hence, there exists a unique solution  $u$  to (2.3) satisfying (2.4) with  $\chi$  given by (2.20).

**Remark 2.6.** Note that perfect plasticity, given by  $k = 0$ , does not satisfy the required conditions as **M1** is violated. Thus we have only considered  $k \in (0, 1)$ .

**Example 2 (Linear/root elasticity).** As a variant of Example 1, we consider the one-dimensional function

$$\sigma(u_x) = \begin{cases} -\frac{1}{2}\sqrt{-u_x}, & u_x < -\frac{1}{4}, \\ u_x, & -\frac{1}{4} \leq u_x \leq \frac{1}{4}, \\ \frac{1}{2}\sqrt{u_x}, & \frac{1}{4} < u_x. \end{cases}$$

Here we have taken  $U_Y = \frac{1}{4}$ . Once again, this satisfies the special case (2.6) with

$$a(r) = \begin{cases} 1, & r \leq \frac{1}{4}, \\ \frac{1}{2\sqrt{r}}, & r > \frac{1}{4}. \end{cases} \quad (2.21)$$

We note that as  $\xi \rightarrow \infty$ , we have  $\sigma'(\xi) \rightarrow 0$ . Because of this, **M1** is not true on all of  $V$ . However, we show that it holds for  $u, v \in V$  such that  $|u_x|, |v_x| \leq M < \infty$ . Let  $\Omega_1, \Omega_2$  be as in (2.17). Then for all  $u, v \in V$  such that  $|u_x|, |v_x| \leq M$  (with  $U_Y = M$ ) we have that

$$\begin{aligned} \sigma'(\xi) &= 1 \quad \text{for } x \in \Omega_1^{u,v}, \\ &= \frac{1}{4\sqrt{|u_x|}} \geq \frac{1}{4\sqrt{M}} \quad \text{for } x \in \Omega_2^{u,v}. \end{aligned}$$

Thus, as in Example 1, for  $u, v \in V$  such that  $|u_x|, |v_x| \leq M < \infty$ ,  $A$  satisfies **M1** with

$$\chi(\|u - v\|_V) = \frac{1}{4\sqrt{M}} \|u - v\|_V. \quad (2.22)$$

As in Remark 2.5, we also see that for such  $u, v$  we have

$$(Au - Av)(u - v) \geq \|u - v\|_{1,2,\Omega_1}^2 + \frac{1}{4\sqrt{M}} \|u - v\|_{1,2,\Omega_2}^2. \quad (2.23)$$

**Condition M2:** As for Example 1, we can establish (2.19) by separately considering the cases  $|u_x| < \frac{1}{4}$  and  $|v_x| < \frac{1}{4}$ ,  $|u_x| \geq \frac{1}{4}$  and  $|v_x| < \frac{1}{4}$ , and  $|u_x| \geq \frac{1}{4}$  and  $|v_x| \geq \frac{1}{4}$ . We then obtain that **M2** is satisfied with  $\Gamma(r) = 1$ . See [3] for details.

**Remark 2.7.** The operator  $A$  in this case is monotone, but not strongly monotone. An alternative theory can still be used in this case to prove existence and uniqueness. However, in this paper, our main focus will be error estimation for cases where the solution is assumed to exist and be regular enough. Condition **M1** will be the most important property.

**Example 3 ( $p$ -Laplacian).** Next, we consider the  $p$ -Laplacian (in one and higher dimensions), given by

$$\sigma(\nabla u) = |\nabla u|^{p-2} \nabla u. \quad (2.24)$$

We will consider the case  $p \geq 2$ . For this, the space  $V$  is given by  $V = W_0^{1,p}(\Omega)$ . Once again this will satisfy the special case (2.6) with

$$a(r) = r^{p-2}.$$

The following theorem taken from [5] shows that  $A$  satisfies **M1** and **M2**.

**Theorem 2.6.** For a given number  $p \geq 2$ , with  $V = W_0^{1,p}(\Omega)$ , let  $A : V \rightarrow V'$  be the operator given by (2.5) and (2.24). Then there exists a constant  $\alpha > 0$ , depending on  $p$ , such that for all  $u, v \in V = W_0^{1,p}(\Omega)$ ,

$$(Au - Av)(u - v) \geq \alpha \|u - v\|_V^p. \quad (2.25)$$

Also, there exists a constant  $M > 0$  such that for all  $u, v \in W_0^{1,p}(\Omega)$ ,

$$\|Au - Av\|_{V'} \leq M(\|u\|_V + \|v\|_V)^{p-2} \|u - v\|_V.$$

For the proof see [5, Theorem 5.3.3].

We see from the above that

$$\begin{aligned} \chi(\|u - v\|_V) &= \alpha \|u - v\|_V^{p-1}, \\ \Gamma(r) &= M(2r)^{p-2}. \end{aligned} \quad (2.26)$$

The proof in [5], only shows the existence of  $\alpha > 0$  without giving a value. However, for our estimators in the sequel, we will need the largest value of  $\alpha$  satisfying (2.26). In [3] it has been established that this optimal value is given by  $\alpha = 1/2^{p-2}$ . This result gives us the following theorem.

**Theorem 2.7.** Let  $A$  be given by (2.5) with  $\sigma(\xi)$  given by (2.24) on the space  $V = W_0^{1,p}(\Omega)$ . Then  $A$  satisfies **M1** with

$$\chi(\|u - v\|_V) = \alpha(p) \|u - v\|_V^{p-1}, \quad (2.27)$$

where  $\alpha(p) = 1/2^{p-2}$ . Moreover, there exists  $M > 0$  such that  $A$  satisfies **M2** with  $\Gamma(r)$  given by (2.26). Hence, there exists a unique solution  $u \in V$  to (2.3) satisfying (2.4) with  $\chi$  given by (2.27).

**Example 4** ( $\varepsilon$ -nonlinearity). As a variant of Example 3 (for the case  $p = 4$ ) we consider operators  $A = A_\varepsilon$  defined for  $\varepsilon > 0$  by

$$\sigma(\nabla u) = \sigma_\varepsilon(\nabla u) = (1 + \varepsilon |\nabla u|^2) \nabla u \quad (2.28)$$

with  $V = W_0^{1,4}(\Omega)$ . We see this case is of form (2.6) with

$$a_\varepsilon(r) = 1 + \varepsilon r^2. \quad (2.29)$$

We choose this example because we can alter the nonlinear effect by changing  $\varepsilon$  (as  $\varepsilon \rightarrow 0$  the nonlinear effect becomes very small).

**M1** and **M2** follow as in Example 3 with

$$\chi_\varepsilon(\xi) = \varepsilon \chi_p(\xi), \quad \Gamma_\varepsilon(r) = 1 + \varepsilon \Gamma_p(r), \quad (2.30)$$

where  $\chi_p$  and  $\Gamma_p$  are given by (2.27) and (2.26). (Note that we can only conclude  $\chi_\varepsilon(\xi) = \varepsilon \chi_p(\xi)$  and not  $\chi_\varepsilon(\xi) = 1 + \varepsilon \chi_p(\xi)$ .)

The presence of  $\varepsilon$  in (2.30) will cause poor results if applied to our estimators later. Hence we use the following result, which shows that the corresponding operator  $A_\varepsilon$  satisfies a version of **M1** uniformly as  $\varepsilon \rightarrow 0$ , but on a space  $H \supset V$  with a weaker norm than  $V$ . The proof is elementary.

**Theorem 2.8.** Let  $A$  given by (2.5) with  $\sigma$  given by (2.28) on the space  $V = W_0^{1,p}$ . Then for any  $\varepsilon > 0$ , and  $F \in V'$ , there exists a unique solution  $u \in V$  to (2.3) with  $\chi$  as in (2.30).

Moreover, let  $H = H_0^1(\Omega) \supset W_0^{1,4}(\Omega) = V$ . Then for any  $u, v \in H$ , the following inequality holds uniformly for  $\varepsilon > 0$  for the above operator  $A_\varepsilon$

$$(Au - Av)(u - v) \geq \chi_H(\|u - v\|_H) \|u - v\|_H, \quad (2.31)$$

with

$$\chi_H(\|u - v\|_H) = \|u - v\|_H. \quad (2.32)$$

**Remark 2.8.** We note that  $H = H_0^1(\Omega)$  is a Hilbert space, which ties in well with the theory developed in the next section. This example will serve as a prototype for other cases where an alternative version of **M1** may be available in a weaker norm. Note that  $A$  will not satisfy **M2** with  $V$  replaced by  $H$ .

### 3. Estimation of the error due to linearization

Nonlinear problems like those described in Section 2 are often solved using a linearized model. Let  $A_L$  be a linear approximation to the nonlinear operator  $A$ . Since  $A_L$  is often discretized by the finite element method (which is usually done in Hilbert spaces) we will assume that there exists a Hilbert space  $H \supset V$  such that  $A_L$  is an operator from  $H \rightarrow H'$  (this implies  $A_L$  is also an operator from  $V \rightarrow V'$ ). The linearized problem is then defined on  $H'$  instead of  $V'$ . Hence we assume  $F \in H'$  and consider the problem

$$A_L u_L = F, \quad F \in H' \quad (3.1)$$

instead of (2.3). We also assume that  $F$  is such that the solution of (3.1) lies in  $V$  rather than just  $H$ . This is in keeping with regularity assumptions typically used in finite element analysis.

To get existence and uniqueness for (3.1), we assume the following:

**L1 :**  $A_L$  satisfies a *coercivity condition* on  $H$ , i.e., there exists  $\alpha_L > 0$  independent of  $u \in H$  such that for all  $u \in H$ ,

$$(A_L u)u \geq \alpha_L \|u\|_H^2.$$

**L2 :**  $A_L$  is bounded on  $H$ , i.e., there exists  $M_L > 0$  independent of  $u$  such that for all  $u \in H$

$$\|A_L u\|_{H'} \leq M_L \|u\|_H.$$

Then we obtain the following theorem.

**Theorem 3.1.** Let  $A_L$  satisfy **L1–L2**. Then for any  $F \in H'$ , there exists a unique solution of (3.1) that satisfies

$$\|u_L\|_H \leq \alpha_L^{-1} \|F\|_{H'}. \quad (3.2)$$

**Remark 3.1.** Conditions **L1–L2** ensure that there exists a bounded, bilinear coercive form  $B(\cdot, \cdot)$  on  $H \times H$  such that  $\forall u, v \in H$ , we have

$$A_L u(v) = B(u, v).$$

In all our examples we take  $H = H_0^1(\Omega)$ . In Examples 1 and 2,  $H = V$  while in Examples 3 and 4,  $V$  is a strict subset of  $H$ . Also, in our examples, we will define  $A_L$  by

$$A_L u_L(v) = \int_{\Omega} \sigma_L(\nabla u_L) \cdot \nabla v \, dx, \quad (3.3)$$

where  $\sigma_L(\xi)$  is a linear approximation of  $\sigma(\xi)$ . Further, we will take

$$\sigma_L(\xi) = a(|\nabla u_0|)\xi, \quad (3.4)$$

where  $u_0$  is some guess for  $u$  (this will moreover give a symmetric bilinear form). For example, for the  $p$ -Laplacian ( $p > 2$ ) we take

$$A_L u_L(v) = \int_{\Omega} |\nabla u_0(x)|^{p-2} \nabla u_L \cdot \nabla v \, dx.$$

For this we then have, with  $H = H_0^1(\Omega)$ , that

$$A_L u(u) = \int_{\Omega} |\nabla u_0(x)|^{p-2} \nabla u \cdot \nabla u \, dx \geq \min_{x \in \Omega} |\nabla u_0(x)|^{p-2} \|u\|_H^2.$$



Thus  $A_L$  satisfies **L1** with  $\alpha_L = \min_{x \in \Omega} |\nabla u_0(x)|^{p-2}$ . Also,

$$\begin{aligned} A_L u(v) &= \int_{\Omega} |\nabla u_0(x)|^{p-2} \nabla u \cdot \nabla v \, dx \leq \max_{x \in \Omega} |\nabla u_0(x)|^{p-2} \int_{\Omega} \nabla u \cdot \nabla v \, dx \\ &\leq \max_{x \in \Omega} |\nabla u_0(x)|^{p-2} \|u\|_H \|v\|_H, \end{aligned}$$

showing  $A_L$  satisfies **L2** with  $M_L = \max_{x \in \Omega} |\nabla u_0(x)|^{p-2}$ .

The same holds true in general—**L1** and **L2** hold with

$$\alpha_L = \min_{x \in \Omega} |a(|\nabla u_0(x)|)|, \quad (3.5)$$

and

$$M_L = \max_{x \in \Omega} |a(|\nabla u_0(x)|)|. \quad (3.6)$$

(Note that the above may not be the optimal values of  $\alpha_L$  and  $M_L$  that can be taken in **L1** and **L2**.)

**Remark 3.2.** In some cases (see Example 3 in Section 3.2 ahead), we may have a  $u_0$  that leads to  $\alpha_L = 0$ . In such cases, we can try to suitably modify  $A_L$  to ensure that  $\alpha_L \neq 0$  (see Remark 3.8).

### 3.1. Theoretical estimates

Our goal in this paper is to devise a method to estimate the modeling error  $e_L = u - u_L$ . To this end, we have the following theorem, which provides a guaranteed upper estimator for this error.

**Theorem 3.2.** Let  $A$  satisfy **M0–M2** and let  $A_L$  be a linear approximation of  $A$  that satisfies **L1–L2**. Also, let  $u$  solve (2.3) and  $u_L$  solve (3.1), with  $F$  such that  $u_L \in V$ . Then

$$\chi(\|e_L\|_V) \leq \|A_L u_L - A u_L\|_{V'}. \quad (3.7)$$

**Proof.** Using **M1** we have

$$\begin{aligned} \chi(\|u - u_L\|_V) &\leq (Au - Au_L) \left( \frac{u - u_L}{\|u - u_L\|_V} \right) \\ &\leq \sup_{\substack{v \in V \\ \|v\|_V=1}} (Au - Au_L)(v) \\ &= \sup_{\substack{v \in V \\ \|v\|_V=1}} (F(v) - Au_L(v)). \end{aligned}$$

Since  $V \subset H$  we have  $A_L u_L(v) = F(v)$  for all  $v \in V \subset H$ , and thus

$$\begin{aligned} \chi(\|u - u_L\|_V) &\leq \sup_{\substack{v \in V \\ \|v\|_V=1}} (A_L u_L - Au_L)(v) \\ &= \|A_L u_L - Au_L\|_{V'}. \quad \square \end{aligned}$$

**Remark 3.3.** Theorem 3.2 holds even when  $A_L$  is any nonlinear approximation to  $A$  such that (3.1) has a unique solution (e.g., when  $A_L$  is a nonlinear operator also satisfying **M0–M2**).

**Remark 3.4.** If  $\chi(\|e_L\|_V) = \alpha \|e_L\|_V$  for  $\alpha \in \mathbb{R}$  with  $\alpha > 0$ , then (3.7) gives the estimate

$$\|e_L\|_V \leq \frac{1}{\alpha} \|A_L u_L - Au_L\|_{V'}.$$

In addition to the upper estimator in (3.7), we also have the following theorem which gives a lower estimator for the error. The result follows easily from **M2** and (3.1).

**Theorem 3.3.** *Let  $A$  satisfy **M0–M2** and let  $A_L$  be a linear approximation of  $A$  satisfying **L1–L2**. Also, let  $u$  solve (2.3) and  $u_L$  solve (3.1) such that  $u_L \in V$  and  $\|u\|_V, \|u_L\|_V \leq r$ . Then*

$$\|A_L u_L - Au_L\|_{V'} \leq \Gamma(r) \|e_L\|_V. \quad (3.8)$$

Thus by Theorems 3.2 and 3.3 we have for  $u, u_L \in B(0; r)$

$$\Gamma(r)^{-1} \|A_L u_L - Au_L\|_{V'} \leq \|e_L\|_V \leq \chi^{-1} (\|A_L u_L - Au_L\|_{V'}).$$

Hence we have both upper and lower estimators based on the quantity  $\|A_L u_L - Au_L\|_{V'}$ .

To be of practical use, an error estimator must be computable. We now show that the term  $\|A_L u_L - Au_L\|_{V'}$  can be bounded above to give a *computable upper estimator*  $\mathcal{E}_V(e_L)$  for  $\|e_L\|_V$ .

Let  $A_L$  be given by (3.3) for some (linear) function  $\sigma_L$ . Then we have for  $V = W^{1,p}(\Omega)$  that

$$\|A_L u_L - Au_L\|_{V'} \leq \|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,q}$$

using the argument of Lemma 2.3. Defining

$$\mathcal{E}_V(e_L) = \chi^{-1} (\|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,q}) \quad (3.9)$$

then gives a computable guaranteed upper estimator satisfying

$$\|e_L\|_V \leq \mathcal{E}_V(e_L).$$

(We have used the fact that since  $\chi$  is increasing  $\chi^{-1}$  must also be increasing.)

**Remark 3.5.** Note that with Dirichlet boundary conditions, we cannot bound  $\|A_L u_L - Au_L\|_{V'}$  below by  $\|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,q}$  to give a similar computable *lower estimator*. Such a bound is, however, possible for some other boundary conditions. For instance, in the case of Neumann boundary conditions, we would have (see Remark 2.4)

$$\|A_L u_L - Au_L\|_{V'} = \|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,q},$$

which would allow us to define a computable lower estimator,

$$\underline{\mathcal{E}}_V(e_L) = \Gamma(r)^{-1} \|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,q}$$

such that

$$\underline{\mathcal{E}}_V(e_L) \leq \|e_L\|_V.$$

Let us now consider the case when in addition to  $V$ , the operator  $A$  is also strongly monotone on the larger space  $H$  over which  $A_L$  is defined, i.e., when (2.31) is satisfied (as in Example 4). In this case, we can obtain a computable guaranteed upper estimator for  $\|e_L\|_H$  as well.

**Theorem 3.4.** *In addition to the conditions of Theorem 3.2, let  $A$  satisfy (2.31). Then*

$$\chi_H(\|e_L\|_H) \leq \|A_L u_L - Au_L\|_{H'}. \quad (3.10)$$

**Proof.** Using (2.31) we have

$$\chi_H(\|u - u_L\|_H) \leq (Au - Au_L) \left( \frac{u - u_L}{\|u - u_L\|_H} \right).$$

Since  $u \in V$  and  $u_L \in V$ , then  $w = (u - u_L)/\|u - u_L\|_H \in V$ . Using this, we have

$$\begin{aligned}\chi_H(\|u - u_L\|_H) &\leq (Au - Au_L)(w) \\ &= F(w) - Au_L(w) \\ &= (A_L u_L - Au_L)(w) \\ &\leq \sup_{\substack{v \in H \\ \|v\|_H=1}} (A_L u_L - Au_L)(v) \\ &= \|A_L u_L - Au_L\|_{H'},\end{aligned}$$

from which (3.10) follows.  $\square$

Recall that in all our examples, we have  $H = H_0^1(\Omega)$ . Using the same idea as for  $\mathcal{E}_V(e_L)$ , we can therefore define

$$\mathcal{E}_H(e_L) = \chi_H^{-1}(\|\sigma_L(\nabla u_L) - \sigma(\nabla u_L)\|_{0,2}) \quad (3.11)$$

to obtain

$$\|e_L\|_H \leq \mathcal{E}_H(e_L).$$

This will be used for Example 4.

### 3.2. Computational results

In this section, we present the results of computational experiments designed to test the accuracy of estimators (3.9) and (3.11). We use the effectivity index  $\kappa$  defined by (1.2), with  $\|e_L\|$  taken to be  $\|e_L\|_V$  or  $\|e_L\|_H$  as specified ahead.

In some cases, like Example 1 and parts of Example 2, it is possible to calculate  $u_L$  exactly. However, when this becomes difficult, we use the finite element method with “sufficient” accuracy to obtain it. We use basis functions of degree  $p$  over  $n$  elements, where  $p$  and  $n$  are chosen to be large enough so that increasing either one will not change  $\|e_L\|_V$  by more than .001%. Thus, the error introduced by the finite element method can be considered negligible and the solution  $u_L$  obtained from this method can be taken to be the required “exact” solution to the linear problem (3.1), sufficiently accurate for our purposes.

**Example 1.** We first consider bilinear elasticity, from Section 2 with  $V = H_0^1(\Omega) = H$  on the domain  $\Omega = (0, 1)$ . We use the function

$$f = 1$$

to define the functional  $F$  by (2.7). It can then be verified that the true solution to the nonlinear problem (2.3) is given by (here  $r = U_Y$ )

$$u = \begin{cases} \frac{1}{k} \left( -\frac{1}{2}x^2 - (1-k)rx + \frac{1}{2}x \right), & 0 < x < -r + \frac{1}{2}, \\ -\frac{1}{2}x^2 + \frac{1}{2}x + \frac{1}{k} \left( \frac{1}{2}(1-k)r^2 + \frac{1}{2}(k-1)r + \frac{1}{8}(1-k) \right), & -r + \frac{1}{2} \leq x \leq r + \frac{1}{2}, \\ \frac{1}{k} \left( -\frac{1}{2}x^2 - (k-1)rx + \frac{1}{2}x - r(1-k) \right), & r + \frac{1}{2} < x < 1. \end{cases} \quad (3.12)$$

We let  $A_L$  be defined by (3.3), (3.4), with  $a(|(u_0)_x|)$  given by (2.16). Then using an initial guess of  $u_0 = 0$ , (2.16) simply gives  $a \equiv 1$ . Thus, for this example  $\alpha_L = M_L = 1$ . The exact solution of the linear problem (3.1) turns out to be

$$u_L = -\frac{1}{2}x^2 + \frac{1}{2}x.$$

From these solutions  $u, u_L$  we see

$$\begin{aligned}\|e_L\|_V &= \left( \int_0^1 |u_x - (u_L)_x|^2 dx \right)^{1/2} \\ &= \left( 2 \left( -\frac{1}{3}r^3 + \frac{1}{2}r^2 - \frac{1}{4}r + \frac{1}{24} \right) + \frac{2}{k} \left( \frac{2}{3}r^3 - r^2 + \frac{1}{2}r - \frac{1}{12} \right) \right. \\ &\quad \left. + \frac{2}{k^2} \left( -\frac{1}{3}r^3 + \frac{1}{2}r^2 - \frac{1}{4}r + \frac{1}{24} \right) \right)^{1/2}.\end{aligned}\quad (3.13)$$

Also, we have  $u_x = (u_L)_x$  on  $[-r + \frac{1}{2}, r + \frac{1}{2}]$  so that

$$\begin{aligned}\mathcal{E}_V(e_L) &= \frac{1}{k} \left( \int_0^1 |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx \right)^{1/2} \\ &= \frac{1}{k} \left( \int_0^{-r+1/2} |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx + \int_{r+1/2}^1 |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx \right)^{1/2} \\ &= \frac{1}{k} \left( 2k^2 \left( -\frac{1}{3}r^3 + \frac{1}{2}r^2 - \frac{1}{4}r + \frac{1}{24} \right) + 2k \left( \frac{2}{3}r^3 - r^2 + \frac{1}{2}r - \frac{1}{12} \right) \right. \\ &\quad \left. + 2 \left( -\frac{1}{3}r^3 + \frac{1}{2}r^2 - \frac{1}{4}r + \frac{1}{24} \right) \right)^{1/2}.\end{aligned}\quad (3.14)$$

From (3.13) and (3.14) we see that for any  $k$  and  $r$ ,

$$\|e_L\|_V = \mathcal{E}_V(e_L).$$

Hence, for this example, our estimator turns out to be *exact*, i.e., we have  $\kappa = 1$ .

**Example 2.** Next, we consider linear/root elasticity with  $V = H_0^1(\Omega) = H$  on the domain  $\Omega = (0, 1)$ . We again use the function

$$f = 1$$

to define the functional  $F$ . It can be shown that the true solution to the nonlinear problem is

$$u = \begin{cases} \frac{4}{3}x^3 - 2x^2 + x, & 0 < x < \frac{1}{4}, \\ -\frac{1}{2}x^2 + \frac{1}{2}x + \frac{5}{96}, & \frac{1}{4} \leq x \leq \frac{3}{4}, \\ -\frac{4}{3}x^3 + 2x^2 - x + \frac{1}{3}, & \frac{3}{4} < x < 1. \end{cases}\quad (3.15)$$

We let  $A_L$  be defined by (3.3), (3.4), (2.21), and use an initial guess of  $u_0 = 0$ , which again gives  $a \equiv 1$ . The exact solution to the linear problem (3.1) is the same as in Example 1,

$$u_L = -\frac{1}{2}x^2 + \frac{1}{2}x.$$

From these solutions  $u, u_L$  we see

$$\begin{aligned}\|e_L\|_V &= \left( \int_0^1 |u_x - (u_L)_x|^2 dx \right)^{1/2} \\ &\approx 0.17970.\end{aligned}$$

Also, we have  $u_x = (u_L)_x$  on  $[\frac{1}{4}, \frac{3}{4}]$  and  $|u_x|, |(u_L)_x| \leq 1$ . Therefore, using (2.22) and  $\alpha_L = M_L = 1$  we have

$$\begin{aligned}\mathcal{E}_V(e_L) &= 4 \left( \int_0^1 |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx \right)^{1/2} \\ &= 4 \left( \int_0^{1/4} |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx + \int_{3/4}^1 |\sigma_L((u_L)_x) - \sigma((u_L)_x)|^2 dx \right)^{1/2} \\ &\approx 0.23222.\end{aligned}$$

Thus, for this example, we have

$$\kappa \approx 1.29226.$$

Next, we use the initial guess  $u_0 = -\frac{1}{2}x^2 + \frac{1}{2}x$ , which is closer to the true solution. We do not calculate  $u_L$  analytically, but instead use a sufficiently accurate finite element method to obtain  $u_L$ . We obtain the error

$$\|e_L\|_V = 0.10275,$$

and again using  $M_L = 1$ , the estimator

$$\mathcal{E}_V(e_L) = 0.12422.$$

Thus, for this initial guess,  $\kappa$  has been improved to

$$\kappa = 1.20895.$$

**Remark 3.6.** In the above example, we knew  $u$  and  $u_L$ , so we were able to accurately take  $M = 1$ . However, in general,  $u$  will not be known, so the best we could do is choose  $M$  to be the maximum of  $|(u_L)_x|$ . Although this might still give a good estimator in computations, we may no longer be able to guarantee that it is an upper estimator.

**Example 3.** Next we consider the  $p$ -Laplacian with  $p = 2.5$  on the domain  $\Omega = (0, 1)$ .

First, we take  $f$  to satisfy the true solution

$$u = x(1 - x).$$

We define the linear operator  $A_L$  given by (3.3) using

$$\sigma_L(\xi) = |(u_0)_x|^{p-2}\xi.$$

We choose a series of initial guesses which converge to the true solution to see how the effectivity index behaves as  $\|e_L\|_V$  decreases. We divide the domain  $\Omega$  into  $M$  equal intervals and take  $u_0$  to be the piecewise linear interpolant of the true solution  $u = x(1 - x)$  for  $M = 2, 4$ , and  $8$ . We obtain  $u_L$  by using the finite element method with  $N = 20$  elements and basis functions of degree less than or equal to  $8$ .

From Table 1 we see that as  $u_0 \rightarrow u$  both the error  $\|e_L\|_V$  and the estimator  $\mathcal{E}_V(e_L)$  decrease. We also see that the effectivity index  $\kappa$  is increasing. This suggests that while the estimator is qualitatively correct, it may not necessarily decrease to 0 as  $\|e_L\|_V \rightarrow 0$  for this example.

Table 1  
True error and estimate for Example 3

$M$	$\ e_L\ _V$	$\mathcal{E}_V(e_L)$	$\kappa$
2	0.17738	0.44583	2.51348
4	0.08329	0.24384	2.92773
8	0.03873	0.14179	3.66115

Table 2  
True error and estimate for Example 3

$M$	$\ e_L\ _V$	$\mathcal{E}_V(e_L)$	$\kappa$
2	0.08618	0.23942	2.77814
4	0.04623	0.14349	3.10376
8	0.02367	0.08654	3.65631

Next, we use the function  $f=1$  to define the functional  $F$ . (This gives more realistic exact solution for such problems.) For any  $p > 2$ , the exact solution of (2.3) is

$$u = -\frac{p-1}{p} \left| x - \frac{1}{2} \right|^{p/(p-1)} + \frac{p-1}{p} \left( \frac{1}{2} \right)^{p/(p-1)}. \quad (3.16)$$

As shown by Exercise 5.3.1 in [5],

$$\begin{aligned} u &\in W^{2,p} \quad \text{if } 1 < p < \frac{3+\sqrt{5}}{2}, \\ u &\notin W^{2,p} \quad \text{if } \frac{3+\sqrt{5}}{2} < p. \end{aligned} \quad (3.17)$$

We again define the linear operator  $A_L$  given by (3.3) using

$$\sigma_L(\xi) = |(u_0)_x|^{p-2} \xi$$

and divide the domain  $\Omega$  into  $M$  intervals and taking  $u_0$  to be the piecewise linear interpolant of  $u$  given by (3.16) for  $M=2, 4$ , and 8. We obtain  $u_L$  by using the finite element method with  $N=20$  elements and basis functions of degree less than or equal to 8.

The behavior of the error and the estimator are shown below for this exact solution. From Table 2 we see that both the error  $\|e_L\|_V$  and the estimator  $\mathcal{E}_V(e_L)$  decrease as  $M$  increases. Once again, the estimator is qualitatively correct, but does not decrease to 0 as  $\|e_L\|_V \rightarrow 0$  for this example. The reasons for this are discussed in the remarks below.

**Remark 3.7.** One factor in the above over-estimation of the true error by  $\mathcal{E}_V$  is the value of  $\alpha(p)$  used in the estimator. The optimal  $\alpha(p)$  used from [3] is calculated using a “worst-case” scenario, and can give a much smaller value of  $\alpha$  in (2.27) than may be typically encountered in a problem. This causes over-estimation.

**Remark 3.8.** For a differentiable  $u_0$  on  $(0, 1)$  which vanishes at 0 and 1, the mean value theorem gives  $(u_0)_x(t) = 0$  for some  $t \in (0, 1)$ , i.e., the coefficient  $a$  vanishes at  $t$ . Hence,  $\alpha_L$  given by (3.5) will be 0 in condition **L1** for such a  $u_0$ . Our choice of  $u_0$  (which is not differentiable) above avoids this problem, since  $\alpha_L > 0$  for the interpolant over  $M$  subintervals. However, the underlying danger of  $\alpha_L$  being close to zero remains for this problem whenever we choose 0 boundary conditions.

**Example 4.** Finally, we consider the  $\varepsilon$ -nonlinearity from Section 2, using  $\mathcal{E}_H(e_L)$  defined by (3.11) to estimate  $\|e_L\|_H$ , with  $H = H_0^1(\Omega)$  and  $V = W_0^{1,4}(\Omega)$ . For this example, recall that the operator  $A$  given by (2.5) uses  $\sigma(u_x)$  given by

$$\sigma(u_x) = (1 + \varepsilon|u_x|^2)u_x.$$

The function  $f$  used to define the functional  $F$  in (2.7) is chosen so that the exact solution is

$$u = (1 - x) \sin(x/k). \quad (3.18)$$

The linear operator  $A_L$  given by (3.3) uses

$$\sigma_L(\xi) = (1 + \varepsilon|(u_0)_x|^2)\xi,$$

where  $u_0$  is an initial guess for  $u$ .

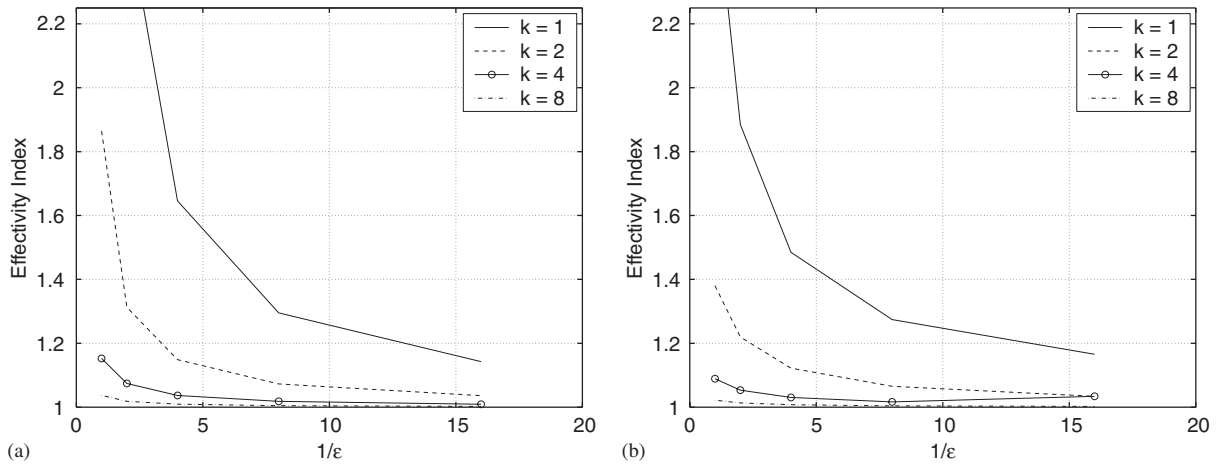


Fig. 1. Effectivity Index,  $\kappa$  for  $u = (1 - x) \sin(x/k)$  and  $\sigma(u_x) = (1 + \varepsilon|u_x|^2)u_x$ : (a) Effectivity index for  $H^1$  semi-norm estimate with  $u_0 = 0$ ; (b) Effectivity index for  $H^1$  semi-norm estimate with  $u_0 = x^2 - x$ .

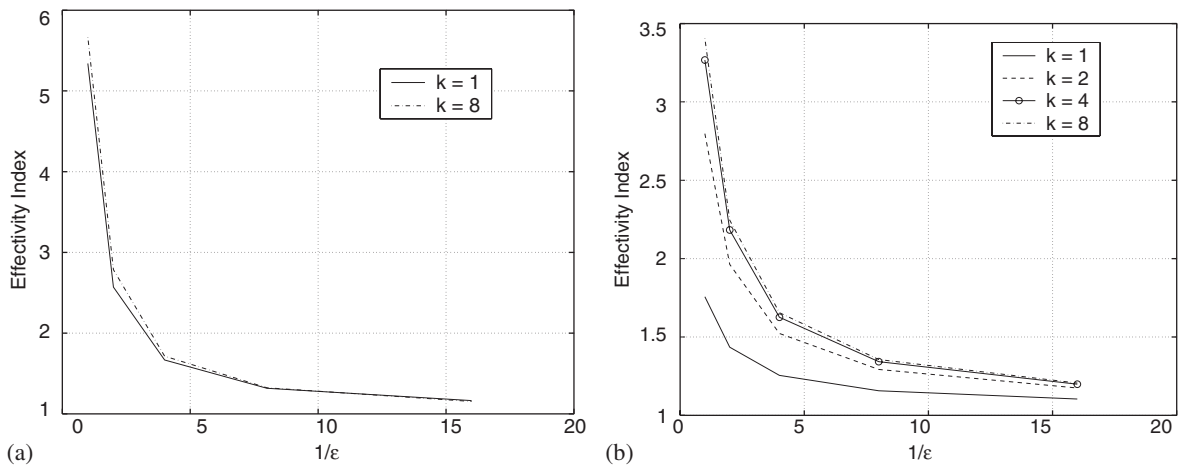


Fig. 2. Effectivity index for  $u = (x^2 - x) \cos(x)$  and  $\sigma(u_x) = (1 + \varepsilon|u_x|^2)u_x$ : (a) Effectivity index for  $H^1$  semi-norm estimate using an initial guess of  $u_0 = 0$ . The graphs for  $k = 2, 4$  are similar; (b) Effectivity index for  $H^1$  semi-norm estimate using an initial guess of  $u_0 = x^2 - x$ .

Fig. 1 shows the effectivity index  $\kappa$  for the true solution  $u = (1 - x) \sin(x/k)$  using an initial guess of both  $u_0 = 0$  and  $u_0 = x^2 - x$ . As expected, we see that as  $\varepsilon \rightarrow 0$ , decreasing the error  $\|e_L\|_H$ , the effectivity index approaches 1. Also, as  $k$  increases,  $u_x$  decreases, again decreasing the effect of the nonlinearity and causing  $\kappa$  to approach 1. Fig. 2 shows the effectivity index for the true solution  $u = (x^2 - x) \cos(x/k)$ . Again we see that as  $\varepsilon \rightarrow 0$  the effectivity index decreases approaching 1. Hence this numerical evidence shows that this estimator decreases to 0 as  $\|e_L\|_H \rightarrow 0$  for the cases considered, a behavior that is similar in spirit to “asymptotic exactness.”

### 3.3. Conclusions

From the above, it is seen that the upper estimator developed in this paper has an effectivity index that behaves very well for some of the examples considered, while for others, it is observed to be in a higher range (between 3 and 4 for Example 3, for instance). As explained in the Introduction, this would still be useful in the context of several nonlinear problems, where “constants” such as yield stress come into play, since these can have a high degree of uncertainty. The corresponding errors induced by such uncertainties in the desired quantities of interest (obtained, for instance, by

a sensitivity analysis) would then be one or several orders higher than the error due to linearization. Hence, one could afford to have a large effectivity index in the linearization estimator.

The computations performed here have been in one dimension, since it is easiest to get an “exact” solution (either analytically or by overkill, using finite elements) in these cases. In [3,4], some computations involving Examples 3 and 4 in two dimensions are also considered, in the context of estimating both the linearization and finite element discretization errors.

## References

- [1] M. Ainsworth, J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley, New York, 2000.
- [2] I. Babuška, T. Strouboulis, *The Finite Element Method and its Reliability*, Oxford University Press, Oxford, 2001.
- [3] A.L. Chaillou, *A posteriori estimation of the linearization and finite element approximation errors for strongly monotone nonlinear operators*, Ph.D. Thesis, University of Maryland, Baltimore County, 2004.
- [4] A.L. Chaillou, M. Suri, *Computable error estimators for the approximation of nonlinear problems by linearized models*, *Comput. Methods Appl. Mech. Engrg.*, to appear.
- [5] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Company, Amsterdam, 1979.
- [6] W. Han, *A posteriori error analysis for linearization of nonlinear elliptic problems and their discretizations*, *Math. Methods Appl. Sci.* 17 (1994) 487–508.
- [7] T.I. Seidman, *A class of nonlinear elliptic problems*, *J. Differential Equation* 60 (1985) 151–173.
- [8] W.A. Wong, R.J. Bucci, R. Stenz, J.B. Conway, *Tensile and strain controlled fatigue data for certain aluminium alloys for applications in the transport industry*, SAE Technical Paper Series no. 870094, International Congress and Exposition, Detroit, 1987.